Summer 2002

An Experimental Investigation of Direct and Indirect Viewing of a Remote Manipulation

Stephen N. Plishka
Embry-Riddle Aeronautical University - Daytona Beach

AN EXPERIMENTAL INVESTIGATION OF DIRECT AND INDIRECT VIEWING OF A REMOTE MANIPULATION

by

Stephen N. Plishka
B.S., Embry-Riddle Aeronautical University, 1992

A Thesis Submitted to the
Department of Human Factors & Systems
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Human Factors & Systems

Embry-Riddle Aeronautical University
Daytona Beach, Florida
Fall 2002

AN EXPERIMENTAL INVESTIGATION OF
DIRECT AND INDIRECT VIEWING OF A
REMOTE MANIPULATION

by

Stephen N. Plishka

This thesis was prepared under the direction of the candidate's thesis committee chair, Steven Hall, Ph.D., Department of Human Factors & Systems, and has been approved by the members of the thesis committee. It was submitted to the Department of Human Factors & Systems and has been accepted in partial fulfillment of the requirements for the degree of Master of Science in Human Factors & Systems.

THESIS COMMITTEE:

Steven Hall, Ph.D., Chair

Christina Frederick-Recascino, Ph.D., Member

Dennis Vincenzi, Ph.D., Member

Sathya Gangadharan, Ph.D., Member

Shawn Doherty, Ph.D., MS HFS Program Coordinator

Fran Greene, Ph.D., Department Chair, Department of Human Factors & Systems

John Watret, Ph.D., Associate Dean of Student Acedemics

ACKNOWLEDGMENTS

The author would like to express special thanks to all the members of the thesis committee for their assistance in the preparation of this thesis. Specifically, I would like to thank Dr. Steve Hall for his guidance on methodology and data analysis; Dr. Christina Frederick-Recascino for her support and enthusiasm for the topic; Dr. Dennis A. Vincenzi for his guidance on the topic; and Dr. Sathya N. Gangadharan for his guidance and support with the robotic equipment.

The author would also like to express special thanks to C.J. Hutto and David Jones for their assistance with the setup and the running of the experiment. Special thanks also goes to Ben Sapte for his assistance with checking grammar within the report.

Finally, sincere appreciation and gratitude to both the author's friends and family, especially Charles and Joan Plishka, for providing the assistance and encouragement needed to complete the task.

ABSTRACT

The purpose of this study was to evaluate direct and indirect viewing of a remote manipulation. With continued exploration of inner and outer space, the ability to directly manipulate objects is lost due to the nature of operating in harsh environments. Remote viewing and operation of equipment is used in such things as the space shuttle manipulator arm, orthoscopic surgery, undersea exploration, and hazardous material management. Most of these operations do not have the luxury of direct viewing. This study will compare the effects of direct viewing vs. indirect viewing from three different viewing distances of 20cm, 60cm and 100cm.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Many tasks may be easier or necessary to do remotely due to harsh, hazardous or inaccessible environments. An example of this is an astronaut's use of manipulator arms to assemble and maintain components of the International Space Station (ISS). Some of these space station components extend beyond the astronaut's field of view and the operator must rely on various means of remote viewing in order to accomplish the task. With the increasing use of manipulator arms in orthoscopic surgery, construction, space exploration, deep sea exploration and hazardous materials management, it would seem prudent to examine how video monitoring can affect the operation of such equipment (Sheridan, 1992). This study will examine how frame of reference within one meter can affect the operation of such equipment.

Problems that can arise in such situations include a lack of stereovision when using a two-dimensional video monitor. Such problems include the inability to accurately measure components while viewing them indirectly and/or the inability of the operator to judge spatial relationships. In addition, some operations such as remote surgery that use stereoscopic displays can cause eyestrain, stemming from a conflict between monoscopic and stereoscopic cues. Some researchers have gone as far as to suggest further investigation into the 10% of the general population that lacks normal binocular vision. Their ability to accurately judge depth solely on monocular cues may make them more suitable for remote operations over the rest of the general population with normal binocular vision (Reinhardt-Rutland, 1996).

Many studies that have investigated direct and indirect viewing conditions of remote operations were conducted with direct viewing conditions beyond 2.5 meters. Subsequently, the conclusions reached were that further research needs to be conducted at distances of under eight feet to take full advantage of stereoscopic vision because, beyond this distance, the brain relies solely on monoscopic cues. For example, while studying teleoperator performance, Massimino and Sheridan (1994) did not find significant differences in mean task times between direct viewing and video viewing, although they did go on to say that the 2.5 meter distance may have been too great to take full advantage of stereo vision and that further research needs to be conducted in this area (Massimino & Sheridan, 1994).

The intention of this study is to compare operator performance, under direct and indirect viewing conditions in close proximity (below 2.5 meters). More specifically, this research will attempt to address the question of whether or not two different viewing conditions make a difference in the performance levels of remote equipment operations that have viewing conditions one meter and below.

Direct versus Indirect viewing

Direct viewing can be defined as looking directly at an operation, while indirect viewing involves observing some task on a two dimensional video monitor. The question to be addressed is that if there is any discrepancy between the two viewing conditions and the performance levels associated with them. The answer to this lies in depth perception and egocentric sense of direction. The average human moves through a physical world made up of three dimensions, of which they perceive themselves to be at the center. Egocentric direction helps people organize their environments and locate objects according to whether they are up, down, left or right, relative to their own position. More specifically, an egocentric point of view is that from an immersed perspective. Conversely, an exocentric point of view is an, "outside in", or bird's eye, perspective.

Depth perception makes it possible to judge distance relatively accurately within the three-dimensional world. It allows for the automatic judgment of distance when doing such things as reaching for a pencil on a desk (absolute distance, observer to object) and parallel parking between two cars (relative distance, object to object). Depth information sources can be divided into two categories, oculomotor cues and visual cues (Sekuler & Blake, 1994).

Oculomotor cues consist of accommodation and convergence, both dealing with muscular contraction of the eyes. Accommodation is the physical action of focusing the lens within the eye, while convergence is the action of both eyes converging to look at an object; the closer the object, the higher the muscle strain. This muscle strain (or lack thereof) serves as a depth cue (Wickins, 1992). An example of accommodation would be reading a newspaper held relatively close to the face. The amount of strain placed on the lens muscles in the eye is greater than if it were further away; consequently, the indication to the brain is that the newspaper is in close proximity. An example of convergence would be a person following the tip of a pencil while moving it toward their nose. The strain caused by this cross-eyed pattern is referred to as convergence.

Visual cues consist of two types, monocular and binocular. Monocular cues consist of interposition, size and perspective. Interposition is when one object is in front of another and partially obscures it. The brain interprets that if the whole object can be seen, it must be closer than the far-away object. For example, if an airplane is sitting in front of a hangar, it partially obscures it. However, if the whole airplane is still visible, the airplane must be between the viewer and the hangar. The size of objects also assists in depth perception. If the size of an object is already known, then the further away it is, the smaller it will appear. For example, to view a sculpture in the distance, it is easier to get an idea of its size when a person of known size is standing beside it (Sekuler & Blake, 1994). Finally, perspective cues are those that refer to appearance of objects as they recede into the distance. For example, railroad tracks appear to converge in the distance when in fact they continue to be parallel.

Binocular cues provide stereovision, accomplished by two eyes viewing an object from slightly different angles. This is referred to as retinal disparity, and applies only at distances below 2.5 to 3.5 meters (depending on the individual). Beyond this, there is essentially no disparity and the brain relies singularly on monocular cues (Sekuler & Blake, 1994). The 2.5 to 3.5 meter limit represents the threshold for stereopsis but does not necessarily mean that it will be an effective visual cue at those distances. Studies have shown that as viewing distances increase beyond one meter, the decreasing convergence angle of the eyes changes very little beyond that point (Boff, Kaufman & Thomas, 1986). The lower the convergence angle, the lower the disparity and the less stereopsis will be effective. Stereopsis is a very effective depth cue at one meter and below but drops off significantly at two meters and beyond (Surdick, Davis, King, Corso, Shapiro, Hodges, & Elliot, 1994).

Therefore, while performing a remote manipulation below 2.5 to 3.5 meters while looking through a window, (direct viewing) all of these cues work together to provide a relatively accurate judgment of depth. However, if an operator uses only a two dimensional video monitor, they will lose oculomotor and binocular cues, and will only be left with monocular cues. As a result, that operator's depth judgment may not be as accurate as it normally would be.

Frame of Reference

With the construction of the International Space Station (ISS) well under way, robotics are being used to aid in its construction and maintenance. Robotic arms and anthropomorphic telerobots are two examples of this. These robotic arms will move along the main truss of the station to manipulate large components and will therefore not always be in direct sight of the operator. Instead, the robotic arms will be remotely operated from inside the space station using closed circuit television. The anthropomorphic telerobots will be able to maneuver all over the station to make repairs via remote control from inside the space station. The telerobots will, more than likely, use two cameras, arranged like "eyes", to provide stereoscopic vision to the remote operator. This camera arrangement places the operator "within" the frame of reference giving them an egocentric point on view. While discussing instrument design, frame of reference has been found to influence spatial awareness – God's eye (exocentric) or pilot's eye (egocentric) (Barfield, Rosenber & Furness, 1995). An example of an exocentric point of view would be an air traffic controller's radarscope, portraying a bird's-eye view of the aircraft being monitored, whereas an egocentric point of view places the operator in the center, such as the viewpoint from a movable security camera.

Humans are inherently egocentric. When asking a participant to locate an object within a room, they may describe it as, 'to my right' or 'to my left'. Any such conclusion is based on one aspect of spatial perception – that the object is somewhere in relation to the viewer. In addition, this awareness determines the motor response to reach for the object or otherwise get closer to it (Harris, 1965).

A major issue is the impact of displaced vision, from its normal orientation in reference to one's body, on frame of reference. For example, studies displacing visual information have been going on for over a century (Cunningham & Welch, 1994). The majority of those studies used prisms to invert, rotate or displace visual information in some manner. The accuracy of hand coordination can be noticeable disturbed by such visual-motor relocation (Cunningham & Welch, 1994). This disruption does not remain in effect during vision displacement; the participants in these experiments adapt after a time and are able to function normally after their perception becomes reoriented to the displacement (Bedford, 1993; Cunningham & Welch, 1994; Harris, 1965; Hay & Pick, 1966; Stratton, 1896).

Adaptation time varies with the type of visual displacement; adaptation to inversion or reversal of the visual field may take many days or weeks, while adaptation to sideways displacement may take just a few minutes (Harris, 1965). Harris (1965) offers five concepts of adaptation that can account for the rapid improvement in reaching for objects seen through prisms:

1. Conscious correction of one's aim.
2. Altered visual perception.
3. Reorientation of the perceptual frame of reference.
4. Visuomotor re-correlation
5. Motor-response learning.

Conscious correction of one's aim occurs when the subject misses the target and corrects for any spatial misjudgment due to altered vision. If the altered vision is subsequently removed, the subject adapts readily to operating normally.

Altered visual perception, what the eye "sees" and the brain "perceives", can be affected by what the brain thinks it should see in terms of spatial positioning. This phenomenon can be demonstrated with or without actual visual alteration such as looking through prisms.

Reorientation of the perceptual frame of reference refers to when perception of any external visual or auditory stimulus is shifted to one side. Perception of the arms, however, is not shifted with the stimulus or else no adaptation would be necessary.

Visuomotor re-correlation happens when visual perception does not change but instead, a visual input is paired with an alternative motor output. For example, inverting glasses flip only those objects that are in the subject's forward viewing space. Items below or to the side of the lens remain unaffected.

Motor-response learning is when new motor skills are acquired in response to a stimulus from a particular spatial location. As an example, learning to reach for a glass while wearing prisms that invert the participant's vision. After a skill is learned, there will still be degraded performance when a different arm movement is used from those previously practiced.

Visual accommodation and adaptation is achieved by putting a large array of non-visual stimuli in contradiction with the optical displacement (Hay & Pick, 1966).

An additional phenomenon is adaptation aftereffect, when subjects need to adapt back to normal vision once optical displacement is removed. This aftereffect suggests that some or all of the adaptation is not under conscious control and can be used to measure the magnitude of adaptation (Cunningham & Welch, 1994).

The use of computers to generate pictures for use as spatial information instruments has been of particular interest in the aerospace industry, but such a frame of reference is not without drawbacks. Just because a system can portray spatial information does not necessarily mean effective spatial information will be transferred to the user (McGreevy & Ellis, 1986). When three-dimensional information is represented on a two-dimensional monitor, the user must interpret and apply what he or she is seeing. Regardless of graphic accuracy, interpretation may result in mental misrepresentation. Adding additional hurdles to such representation is the fact that perspective dictates a large portion of perception (McGreevy & Ellis, 1986).
McGreevy and Ellis also determined that the field of view can affect the perspective in regard to target elevation estimation.

Virtual environments are another area where frames of reference have recently been studied (McCormick, Wickens, Banks, & Yeh, 1998). McCormick et al. (1998, p. 444) defined three generic frames of reference (FOR):

1. Egocentric: an immersed perspective, presenting the user the viewpoint of being inside the virtual environment.
2. Exocentric: an outside in (or bird's-eye) perspective that gives the user the vantage point of overlooking all or a large portion of the virtual environment.
3. Tethered: attempts to incorporate some principles of both exocentric and egocentric frames of reference. The tethered perspective positions the

> user's vantage point behind the representative icon, allowing a wider view of the virtual environment but not to the same extent as the exocentric display. Furthermore, the tethered perspective continues to follow the representative icon as it turns and travels through the environment, thus allowing the user a view of the environment that requires no mental rotations to interpolate position or attitude.

An interesting aspect to this research with respect to remote manipulation is the travel subtask. This requires the user to navigate quickly and accurately throughout the virtual environment to capture or designate an object. The results from this portion of the experiment indicate that the egocentric and tethered frame of reference displays allowed a shorter travel time and require fewer control inputs than exocentric displays. Using exocentricity was not efficient when operating in close proximity to the target object (McCormick et al., 1998). With a remote manipulator arm, the operator will need a high degree of near-target performance. A tethered FOR seems to be a desirable view to use for a remote manipulator arm, providing good qualities for traveling within an environment and a wider field of view than an egocentric one. With a tethered view, a remote manipulator arm could possible be perceived as an extension of the operator.

Closed Circuit Video Monitoring

The Manipulator Flight Demonstration (MFD) was successfully conducted aboard the Space Shuttle in 1997 (Wakabayshi, Matsumoto, Horikawa & Nagatomo, 1998). In order to provide visual information to the crew, the MFD used two video camera sets with one attached to the wrist.

The crew, in the aft flight deck, can select two out of three cameras located in the payload bay and display the views on Closed Circuit TV (CCTV) monitors. They also control the robot arm from this workstation using hand controllers and camera views provided by the CCTV monitors (Wakabayshi et al., 1998).

In a different example of closed circuit video monitoring, endoscopic surgery deals with monitoring in very close relative proximity, within a few millimeters as opposed to the length of a Shuttle cargo bay. Holden and Flach (1996) indicated that the surgical instruments and camera in effect become supplements to the surgeon's perceptual-motor system. However, endoscopic tasks are not without their perceptual drawbacks, endoscopes provide only a narrow, monocular field of view forcing surgeons to probe inside the patient with their instruments to determine relative depths (Holden & Flach, 1996). This monocular field of view would also be a problem for an astronaut using only closed circuit television as a visual reference while using a remote manipulator arm.

In the discussion of this endoscopic simulation, Holden and Flach (1996) noted that participants quickly became accustomed to sensory rearrangement produced by camera and display. When an additional rearrangement took place due to stimuli reversal, skill levels retarded to levels comparable to those at the beginning of training. However, subjects' skills quickly returned with practice.

Combination of Direct viewing and Video Monitoring

Researchers in Japan have been developing and testing the Japanese Experiment Module Remote Manipulator System (JEMRMS) for future space missions. Researchers observed during testing that the control of the main arm through only video monitoring was difficult for large transfer movements by the operator. To improve positioning capability, the system employed pre-programmed operation positioning, controlled by the operators through the use of pre-programmed trajectories that allow the operator to monitor the operation through a window or video monitor. This automation is used as much as practical to save time and reduce crew fatigue. In the manual mode, the operator experienced some difficulties in trying to perceive the distance to a target and accomplish fine positional movements of the payload (Kuraoka et al., 1990).

While performing remote tasks, frame of reference plays a very important role in accomplishing the task effectively and efficiently. Selecting the frame of reference most compatible with the operator's mental model can lead to enhanced situational awareness and thus better operator performance and decision-making with less cognitive workload (Barfield, Rosenburg & Furness, 1995). If a compatible frame of reference is used, the task becomes more intuitive for the operator and more of an extension of the operator's abilities rather than a hindrance.

Remote Equipment Manipulation

Remote manipulation of equipment can be very critical. The endoscopic surgery study by Holden & Flach (1996) showed that moving the camera in relation to the instruments can be very disruptive to coordination.

> The surgeons have told us that camera movements can be very disorienting and on several occasions we have witnessed sharp exchanges from the surgeon when the camera was moved by the assistant during dissection. Also, the surgeons can be disoriented by changing positions relative to the camera. Surgeons talk about having to "turn their brain

> around" in order to adjust to the new perspective. Thus, it is good policy to keep the dynamic relation between camera and instruments as consistent as possible during surgery for good coordination (Holden & Flach, 1996, p. 20).

This is also important for the remote manipulator arms for the ISS since they will be able to move to different positions on the main truss assembly. The remote manipulator arms for the ISS are going to be much more dynamic compared to the remote manipulator arm on the space shuttle.

There have been several studies developing remote manipulator systems and concepts. Some of the things that have been looked at include using the ISS remote arms to capture the orbiter, controlling ISS remote arms from the ground, and dexterous robot arms (Bains, Price & Walter, 1987; Kuraoka, Goma, Shinomiya & Nishida, 1990; McCain, Andary & Hewitt, 1990; Sallanberger, 1997; Wakabayashi, Matsumoto, Horikawa & Nagatoma, 1998).

Direct viewing vs. Video monitoring

While conducting remote operations, vision is the most vital sensory input. Therefore, logic dictates that a direct view of remote equipment operations would be better than using a video monitor. Unfortunately, conditions do not always provide a close or even a direct view of a remote operation. As a result, video monitoring must be employed (Sheridan, 1992). However, there are drawbacks to using video monitoring. Conventional video monitors are monoscopic, causing the operator to lose all stereoscopic cues.

Theory tells us that there should be degradation in performance when transitioning from a direct stereoscopic view to an indirect monoscopic video monitor. Some studies support this theory, and state that stereoscopic viewing is superior to monoscopic viewing (Pepper & Hightower, 1984; Drascic, 1991). However, there are other studies that state that there is little or no difference in performance between monoscopic and stereoscopic viewing (Crooks, Freedman, & Coan, 1975; Hankins & Mixon, 1986; Massimino & Sheridan, 1994; Park & Woldstad, 2000). These studies looked mainly at task execution times and the number of errors committed.

Some problems arise while looking at those studies that did not find any differences between these viewing conditions. For example, a few studies have the viewing distances at 2.5 meters and over. Theory states that while directly viewing objects, stereoscopic cues are effective only below 2.5 meters. These studies even suggested that further research needed to be conducted with viewing distances under eight feet, in order to take full advantage of stereoscopic cues. One study trained the participants with direct viewing. When they ran the experiment they compared monoscopic and stereoscopic video monitors (Crooks, Freedman, & Coan, 1975). However, they failed to compare direct viewing along with the video monitors. Another study used a viewing point in front of the robot and oriented the control to that position. Not only did this give the participant an exocentric point of view but also reverse-mapped the robot arm to the participants own, creating some confusion. Finally, another study used only four participants and dropped the data for one of them when they analyzed it (Hankins & Mixon, 1986).

This study will use some of the recommendations and things that were overlooked from these previous studies. It has been suggested that viewing conditions need to be tested below 2.5 meters of viewing distance (Massimino & Sheridan, 1994). This study will compare participants with a direct view and indirect view of a remote operation at viewing distances one meter and below. It will also measure the accuracy of the task they will perform under each condition.

Statement of Hypothesis

Performance of a specific task using a robotic arm will vary, depending on viewing distance and condition. For direct viewing conditions of one meter or less, it is expected that participant performance will decrease as viewing distance increases. However, for indirect viewing using a video monitor, it is expected that participant performance will remain constant, regardless of viewing distance. Overall, it is expected that performance from participants under direct viewing conditions will exceed that of those under indirect viewing conditions.

Method

Participants

Participants for this study were selected from an Introduction to Psychology class at Embry-Riddle Aeronautical University. A total of 90 people, 72 males and 18 females, participated, ranging in age from 18 to 33. Each participant had normal, or correctable to normal, vision. Participants were also prescreened for stereo-acuity using the Howard-Dolman test, which utilizes the relative distance of two pegs in free space. The participants viewed two black pegs against a white background through a square opening in a long wooden box. They were then instructed to move the pegs back and forth with strings until the pegs appeared to be beside one another. One participant was replaced after failing the stereo-acuity test.

Apparatus

Six major pieces of equipment were used for this study (see Appendix A):

- Sony Video Camera Model no. CCD-TR87
- 15 inch Panasonic color monitor model no. CT13R14V
- Two cubicle style movable walls
- Foam board with 30.3cm by 21.5 cm Plexiglas window
- Chinrest constructed from PVC pipe
- Questech Robot Manipulator Arm Model TCM

Design

This study is a 2(view) x 3(viewing distance) design. The independent variables are type of view (direct vs. indirect viewing) and viewing distance (20cm, 60cm and 100cm). The dependent variable is the accuracy of the completed task, as measured by the number of times the participant drops a ring completely to the bottom of a dowel.

Procedure

This experiment was conducted as a between participants study. Each participant was shown the experimental apparatus with the remote manipulator arm holding the ring 1.25cm directly over the dowel. They were then instructed in the use of the remote manipulator arm and provided with a diagram labeling each manipulator arm joint with the corresponding buttons on the robot arm controller. The top of the dowel was measured to be 119cm from the floor. A chinrest was used to ensure that all direct view participants were at eye level with the top of the dowel, and to keep them from moving their heads to gain different perspectives. The chinrest was also used to provide the appropriate viewing distance, by lining up the center of each leg with the appropriate viewing distance marked on the floor for each direct viewing distance. The 60cm and 100cm viewing conditions used a Plexiglas window to approximate the field of view produced by the video camera and television monitor. The 20cm condition did not use the window for two reasons; first, it was not practical for such a close viewing distance. Secondly, at 20cm, the potential for a difference in the field of view was not an issue between the direct and video conditions. For the indirect view, the window was removed and the video camera lens was leveled with the top of the dowel, and at the appropriate distance, to approximate the same field of view for each direct viewing distance. Then, the right moveable cubicle wall was moved to block the view for the indirect participants.

Upon the completion of the instructions for the remote manipulator arm, the participants were asked to drop a thick flat-sided ring – the center plastic insert from a roll of cellophane tape – over a dowel, placed in the arm's work area. Participants were asked to turn around while the arm was set to a start position with the ring in the arm's gripper. Accuracy was judged by the amount of times that the ring fell to the bottom of the dowel out of ten trials. The manipulator arm was reset to the same start position with the ring for each trial. Rings that were hung up on the end of the dowel, or rings that missed the dowel, were considered errors. This procedure was conducted for each viewing condition and distance.

Results

The purpose of this study was to examine the relationship between participant distance from the target object, and the viewing mode used by the participant The success rates with a robotically manipulated task were then noted The success rate was measured as the number of times out of ten trials, that a participant was able to drop a ring onto a dowel Table 1 summarizes performance data for each of the experimental groups Figure 1 illustrates the means for the two types of view at 20, 60, and 100cm

Table 1

Group Means and Standard Deviations

View Type	Distance	*M*	*SD*	*n*
Direct	20 cm	8 87	1 96	15
	60 cm	8 40	2 44	15
	100 cm	7 73	2 25	15
	Total	8 33	2 22	45
Indirect	20 cm	2 80	1 82	15
	60 cm	2 27	1 58	15
	100 cm	0 67	0 98	15
	Total	1 91	1 73	45
Total	20 cm	5 83	3 60	30
	60 cm	5 33	3 72	30
	100 cm	4 20	3 98	30
	Total	5 12	3 79	90

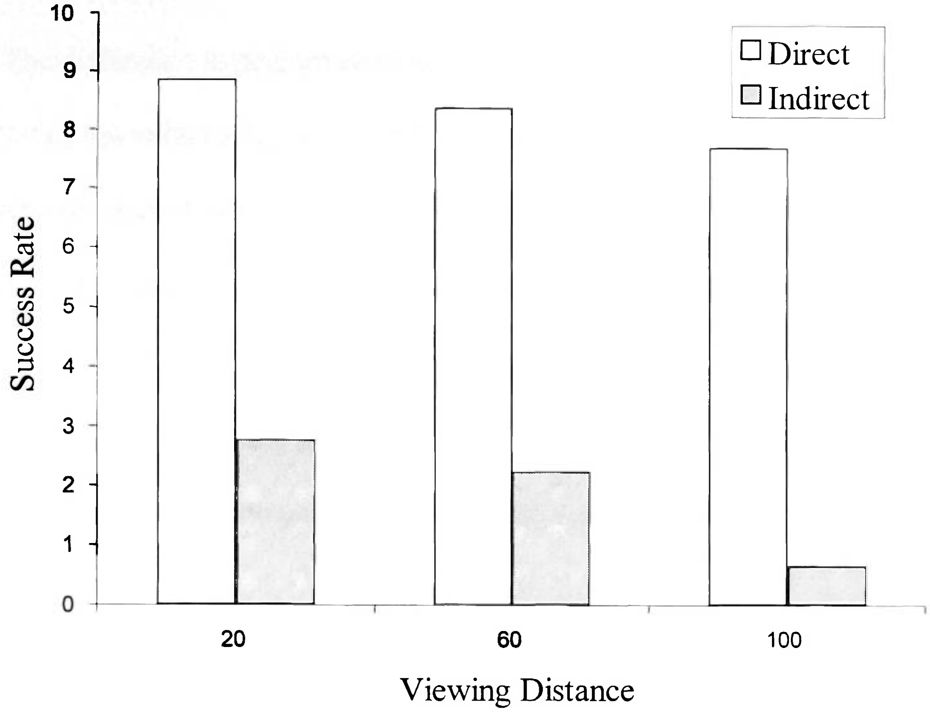

Figure 1. Group means for the two types of view at 20, 60, and 100cm.

The data was analyzed using a between subjects factorial ANOVA, with view type and viewing distance as factors. An alpha level of .05 was used for all significance testing. The results of the ANOVA indicated a significant main effect for both view type, $F(1, 84) = 257.3$, $p < .001$, and viewing distance, $F(2, 84) = 5.8$, $p = .004$. The interaction term failed to be statistically significant, $F(2, 84) = 0.65$, $p = .53$. Table 2 presents additional information regarding the ANOVA results, including estimates of effect size and power.

Table 2

ANOVA Source Table

Source	*df*	*SS*	*MS*	*F*	*p*	Partial η^2	Power
View Type	1	928.0	928.0	257.3	<.001	.754	1.00
Distance	2	42.0	21.0	5.8	.004	.122	.86
Interaction	2	4.7	2.3	0.7	.525	.015	.16
Error	84	302.9	3.6				
Total	89	1277.7					

The difference in performance as a function of viewing condition (direct vs. indirect) was 6.4 units (SD_{pooled} = 1.98). Follow-up tests on the viewing distance main effect were conducted using Tukey HSD to control for Type I error inflation. All possible pairwise comparisons for the distance groups were made, and the results indicate a significant difference of 1.63 units (p = .004) between the 20 cm and the 100 cm conditions. None of the other group mean differences were statistically significant. Table 3 presents group mean difference data along with adjusted confidence intervals for the mean differences.

Table 3

Group Mean Differences for the Viewing Difference Variable

Group I	Group J	Mean Difference (I – J)	*SE*	*p*	95% Confidence Interval For the Mean Difference Lower Bound	Upper Bound
20 cm	60 cm	0 50	49	567	- 67	1 67
20 cm	100 cm	1 63	49	004	46	2 80
60 cm	100 cm	1 13	49	060	- 04	2 30

Discussion

The purpose of this study was to evaluate the impact of direct and indirect viewing conditions on the performance of a remote manipulation task. With exploration of inner and outer space, the ability to directly manipulate objects is severely restricted, due to the nature of operating in harsh or confined environments. Remote viewing and operation of equipment has application in space exploration, orthoscopic surgery, undersea exploration and hazardous material management. Most of these operations do not have the luxury of direct viewing with a normal frame of reference.

More specifically, the purpose of this study was to evaluate the performance between direct and indirect viewing of a remote manipulation at 20, 60, and 100 centimeters. In the sample of 90 college students, there was a significant difference in performance of a remote operation across viewing conditions. That is to say that those participants completing the manipulation task under the indirect viewing condition performed worse than those under the direct viewing condition. In terms of viewing distance, only the 20cm and 100cm conditions were found to produce significant mean performance differences. The viewing and distance conditions did not produce a significant interaction.

While both the viewing condition factor and the distance factor resulted in performance differences, the magnitude of effect produced by the viewing condition factor was quite large, Cohen's $d = 3.24$, indicating that participants in the direct viewing condition performed over three standard deviations better than those in the indirect viewing condition. This very large effect size could indicate that people rely heavily on stereoscopic vision at a viewing distance of one meter or less when performing a remote task. This also suggests that the impact of viewing format should be seriously considered when designing remotely operated systems. The viewing distance factor also produced significant differences in performance, but the impact of viewing distance on performance was smaller than that produced by the viewing condition factor, Cohen's d = 0.43. Even though this effect is smaller, it is still considered to be of medium size (Hinkle, Wiersma, & Jurs, 1998). This medium effect size implies that the viewing distances of a remote task is also an important factor to consider when designing such tasks, but not as important as viewing mode.

The results indicate a large main effect for the viewing condition, but these results were not exactly in line with the hypothesized interaction between the viewing condition variable and the distance variable. It was hypothesized that performance for the direct view participants would drop off with the increase in viewing distance due to a decrease in retinal disparity. It was also hypothesized that the performance of the indirect view participants would be low and remain relatively constant across all three viewing distances, due to a lack of retinal disparity from viewing the task with a video monitor. This is not what happened; performance for the indirect view participants also dropped off with the increase in viewing distance.

The main reason for this appears to be the lack of binocular cues when using an indirect view, such as a monoscopic video monitor. However, the lack of an interaction between the direct and indirect views suggests that participants may be using monocular cues more than expected. Conversely, it is possible that the lack of interaction could also be attributed to the constant resolution of the camera and video combination with an increase in viewing distance. In other words, as the distance between the object and camera increases, the clarity of the object decreases on the monitor.

The main effect for the distance condition could possibly be explained by two things. The first of these is a decrease in retinal disparity, degrading the performance of the direct view participants. Secondly, as stated above, the constant resolution of the camera and video equipment could cause a reduction in the monitor view clarity over increasing camera distances.

While running participants through this study, some interesting observations were noted. During the first few ring drops in the video viewing condition, the participants usually dropped the ring about 30cm (+/- 10cm) behind the dowel, even after presenting them with the exact video view of the ring being held directly over the dowel. A possible explanation for this is that the robot arm start position was to the right of the dowel and slightly behind. As a result, participants were usually raising the robot arm and approaching the dowel from behind. In addition to this, the participants did not take note of the relative size of the ring and robot arm in relation to the dowel when they were presented with the exact video view of the ring being held directly over the dowel. Participants would also move their heads from side to side while looking at the video monitor as if they were getting a better viewing angle.

Some limitations of this study included the contrast between the color of the ring and the background. Several video viewing participants complained that they had a difficult time seeing the white ring against the off-white background, even though it was being held by the black gripper of the robot arm. Another limitation could involve the task itself. Since the direct view had such a high number of "hits" and the indirect view had such a low number, this indicated that there may perhaps be a task better suited to this type of study. In addition, a lack of training for these conditions could also help explain such a performance disparity. One more possibility could have been not exceeding 100cm of viewing distance and how that might have compared to viewing distances below 100cm.

In order to overcome some of these limitations, there are several potential directions that future research could take in studies similar to this one. A possibility may be to explore why most participants drop the ring behind the dowel in the video condition. Multiple cameras and views may be used to explore this issue. Another possible research direction is changing the color of the ring or the background, to increase visual contrast. Another possibility to reduce potential performance differences in viewing conditions could be remedied with additional task training prior to running the experiment. One more direction for future research could be viewing distances that exceed 100cm. Research of remote operations with viewing distances from 1 meter to 2.5 meters should be conducted to see how the performance compares to that of viewing distances of 1 meter and below.

For the most part, the results of the study substantiated part of the original hypothesis. That is to say that it would appear that there are differences between indirect and direct viewing conditions, at viewing distances of one meter or less. However, the study showed, in contradiction of expectation, that performance under the indirect viewing condition did not stay constant over increasing viewing distances, as predicted.

References

Bains, E., Price, C., & Walter, L. (1987). *Track and Capture of the Orbiter with the Space Station Remote Manipulator System.* (NASA-TM-89221).

Barfield, W., Rosenber, C., & Furness T. (1995). Situation Awareness as a Function of Frame of Reference, Computer-Graphics Eyepoint Elevation, and Geometric Field of View. *The International Journal of Aviation Psychology, 5*(3), 233-256.

Bedford, F. (1993). Perceptual and Cognitive Spatial Learning. *Journal of Experimental Psychology, 19*(3), 517-530.

Boff, K., Kaufman, L., & Thomas, J. (1986). *Handbook of Perception and Human Performance, Volume I, Sensory Process and Perception*. New York : John Wiley and Sons

Crooks, W., Freedman, L., & Coan, P. (1975). Television Systems for Remote Manipulation. *Proceedings Human Factors Society 19th Annual Meeting,* 428-435.

Cunningham, H., & Welch, R. (1994). Multiple Concurrent Visual – Motor Mapping: Implications for Models of Adaptation. *Journal of Experimental Psychology, Human Perception and Performance, 20*(5), 987-999.

Drascic, D. (1991). Skill Acquisition and Task Performance in Teleoperation using Monoscopic and Stereoscopic Video Remote Viewing. *Proceedings of the Human Factors Society 35th Annual Meeting – 1991.* 1367-1371.

Hankins III, W & Mixon, R. (1986). *Baseline Experiments in Teleoperator Control.* (NASA-TP-2547).

Harris, S. (1965). Perceptual Adaptation to Inverted, Reversed, and Displaced Vision. *Psychological Review, 72*(6), 419-444.

Hay, J., & Pick, H.(1966). Visual and Proprioceptive Adaptation to Optical Displacement of the Visual Stimulus. *Journal of Experimental Psychology, 71*(1), 150-158.

Hinkle, D.E., Wiersma, W., & Jurs, S.G. (1998). *Applied Statistics for the Behavioral Sciences*. New York: Houghton-Mifflin.

Holden, J., & Flach, J. (1996). *Perceptual – Motor Coordination in an Endosocpic Surgery Simulation.* Dayton, OH: Wright State University, Psychology Department.

Kuraoka, K., Goma, K., Shinomiya, Y., & Nishida, S. (1990). JEMRMS Operational Performance Verification Approach. *41st Congress of the International Astronautical Federation.*

Massimino, M. & Sheridan, T. (1994). Teleoperator Performance with Varying Force and Visual Feedback. *Human Factors, 36*(1), 145-157

McCain, H., Andary, J., & Hewitt, D. (1990). The Space Station Freedom Flight Telerobotic Servicer: The Design and Evolution of a Dexterous Space Robot. *41st Congress of the International Astronautical Federation.*

McCormick, Wickens, C., Banks, R., & Yeh, M. (1998). Frame of Reference Effects on Scientific Visualization Subtasks. *Human Factors 40*(3), 443-451.

McGreevey, M., & Ellis, S. (1986). The Effect of Perspective Geometry on Judged Direction in Spatial Information Instruments. *Human Factors, 28*(4), 439-456.

Park, S. & Woldstad, J. (2000). Multiple Two-Dimensional Displays as an Alternative to Three-Dimensional Displays in Telerobotic Tasks. *Human Factors 42*(4), 592-603.

Pepper, R. & Hightower, J. (1984). Research Issues in Teleoperator Systems. *Proceedings of the Human Factors Society – 28th Annual Meeting –1984,* 803-807.

Reinhardt-Rutland, A. (1996). Remote Operation: A Selective Review of Research Into Visual Depth Perception. *The Journal of General Psychology, 1996, 123*(3), 237-248.

Sallanberger, C. (1997). Canadian Space Robotic Activities. *48th International Astronautical Congress.*

Sekuler, R. & Blake, R., (1994). *Perception* (3rd ed.). New York: McGraw-Hill

Sheridan, T. (1992). *Telerobotics, Automation, and Human Supervisory Control.* Cambridge, MA: The MIT Press.

Stratton, G. (1896). Some preliminary experiments on vision without inversion of the retinal image. *Psychological Review 3,* 611-617.

Surdick, R., Davis, E., King, R., Corso, G., Shapiro, A., Hodges, L., & Elliot, K. (1994). Relevant Cues for the Visual Perception of Depth: Is Where You See It Where It Is? *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting – 1994,* 1305-1309.

Wakabayashi, S., Matsumoto, K., Horikawa, Y., & Nagatoma, M. (1998). On the Results of the Manipulator Flight Demonstration for the JEM. *48th International Astronautical Congress.*

Wickens, C. (1992). *Engineering Psychology and Human Performance* (2nd ed.). New York: HarperCollins.

APPENDIX

PHOTOGRAPHS OF APPARATUS SETUP

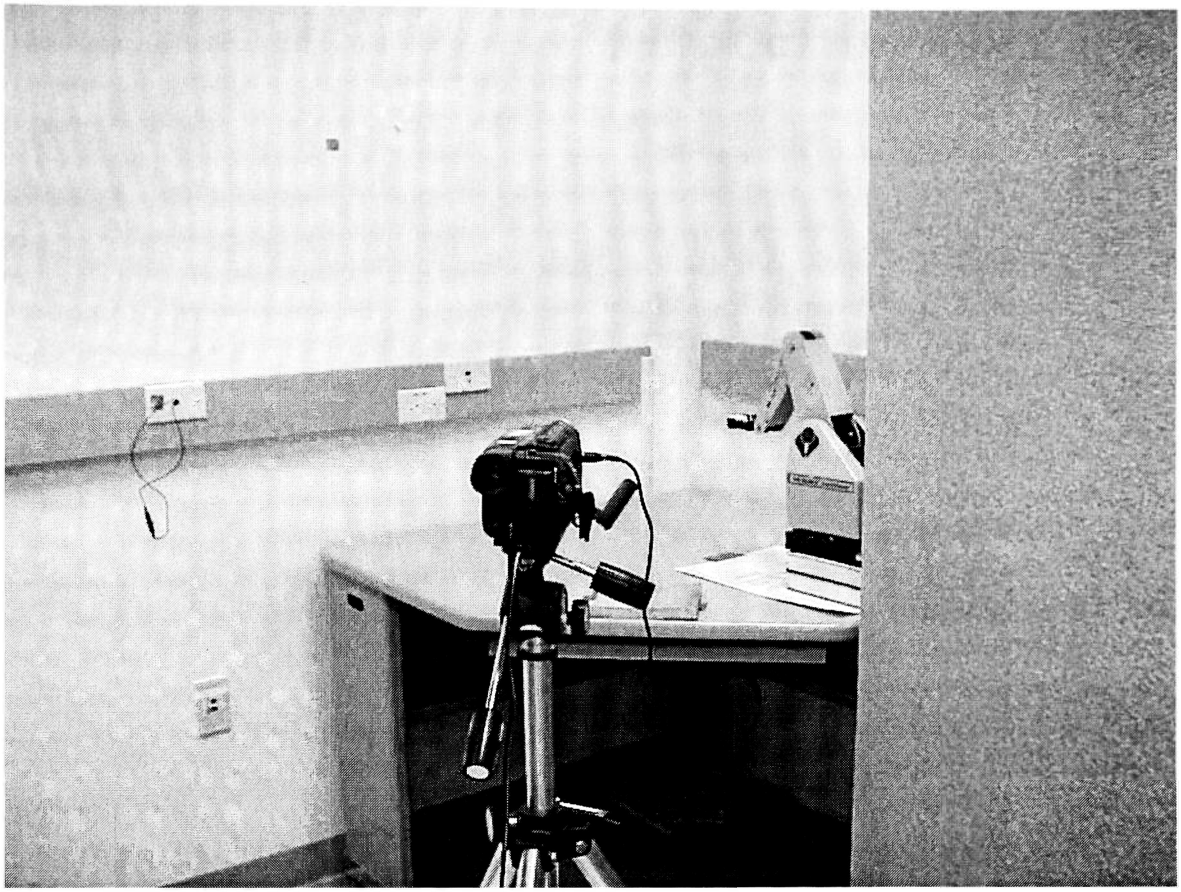

Figure A1. Camera setup for indirect viewing conditions.

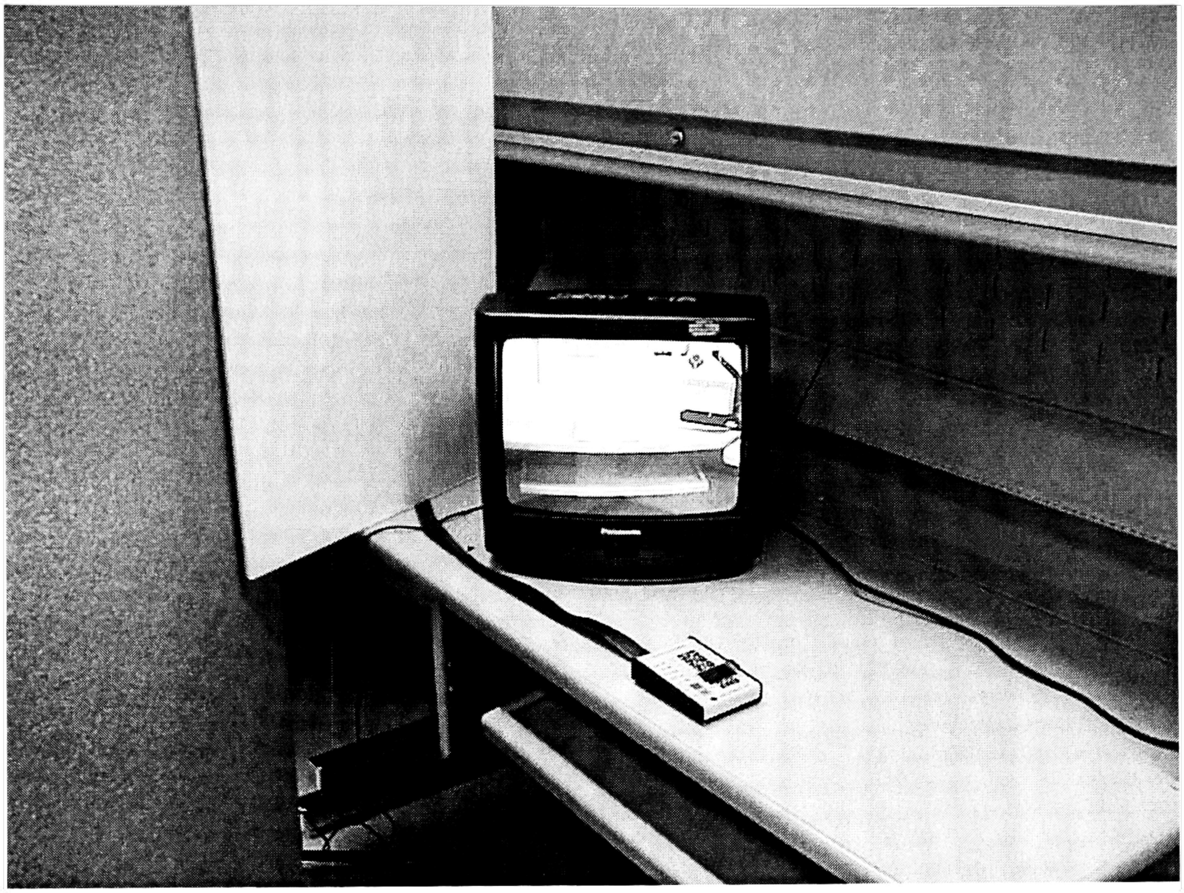

Figure A2. Video viewing station for the indirect viewing conditions.

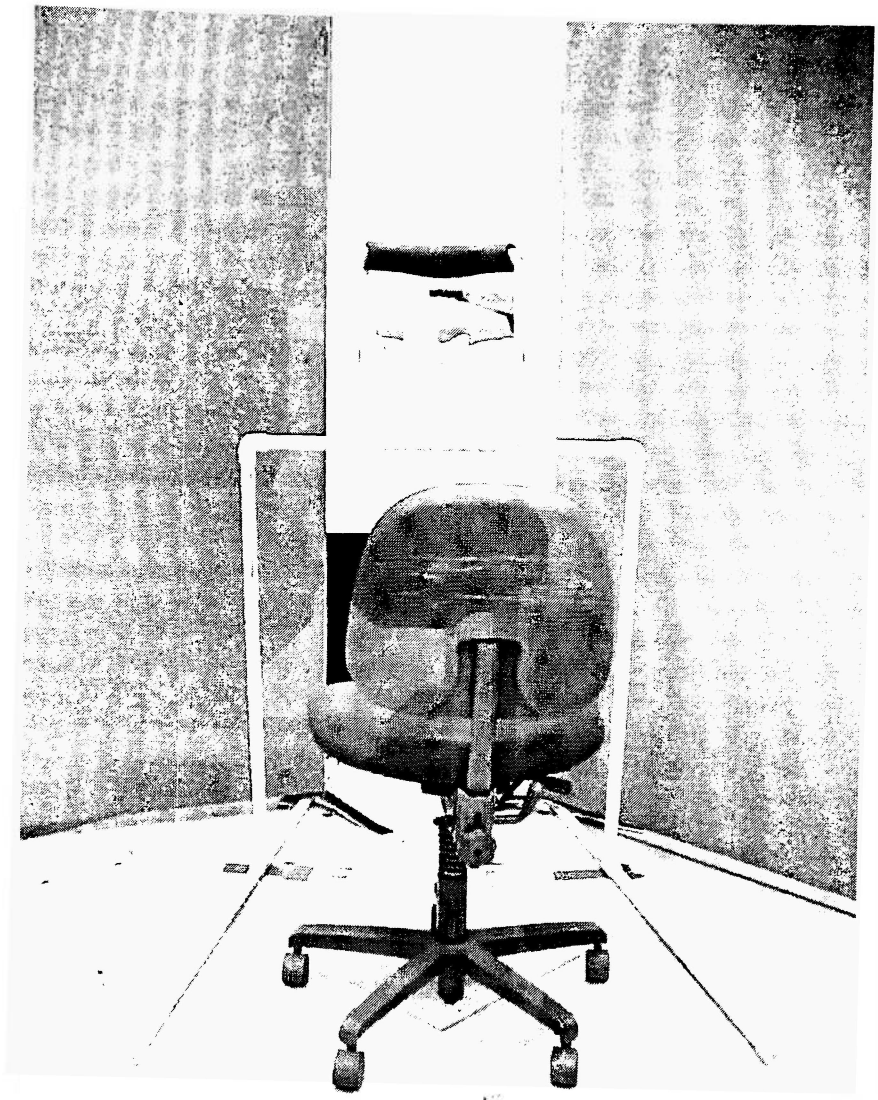

Figure A3. Direct viewing station for direct viewing conditions.